Semi-Automated Techniques for Training Data Extraction for Supervised Classification of Digital Satellite Imagery

A *Thesis* submitted
in Partial Fulfillment of the Requirements
for the Degree of
Master of Technology

by
Ravi Gupta

to the
Department of Civil Engineering
Indian Institute of Technology, Kanpur
August, 1995

CERTIFICATE

Certified that the work contained in the thesis entitled "*Semi-Automated Techniques for Training Data Extraction for Supervised Classification of Digital Satellite Imagery*", by "**Ravi Gupta**", has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

(Onkar Dikshit)
Department of Civil Engineering,
Indian Institute of Technology,
Kanpur.

August, 1995

ABSTRACT

Training area selection and refinement is considered to be an important part of the supervised classification procedure, as it affects the overall classification accuracy of the classifier. Obtaining a representative class accurately and efficiently has been the major goal of this study, which has been attempted by the refinement of traditional data and by applying alternate methods of training area selection..

A novel technique, termed the seed pixel method, of training area delineation was attempted as a means for improving both the efficiency and consistency of training field extraction for images with high resolution and high spectral variability. It was designed to increase the degree of automation of training field extraction, thereby significantly reducing analyst time requirements, while maintaining acceptable levels of accuracy. It is viewed as a potential replacement to other non-traditional image analysis techniques. This algorithm gave 100% training area accuracies and better test area accuracies in heterogeneous classes as compared to the traditional approach.

The basic approach used in the thesis was to apply two levels of classification strategies, one by classifying the images by selecting training areas by conventional methods and by seed pixel method. And the second, by classifying the images after refinement suggested by Buttner et al. (1988) and Mather (1987). These refinement procedures have shown to improve the accuracies of classification.

ACKNOWLEDGEMENT

I am thankful to my guide Dr. Onkar Dikshit for introducing me to the world of Remote Sensing. It was a great experience to have worked under his guidance.

I am grateful to Prof. K. V. G. K. Gokahle and Dr. Nitin Tripathi for the help and encouragement provided inside and outside the classroom.

Thanks to all my friends who helped me in completing this task.

Faults and flaws are all mine.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

Chapter 1

INTRODUCTION

1.1 Background of the Thesis

Automatic classification of the pixels making up a remotely-sensed image involves associating each pixels in the image with a label describing a real world object. It is a problem of recognition in which the numerical values associated with each pixel are required to be identified in terms of observable geographical, geological or other earth surface cover type. This labeling operation, when carried out for all pixels in the area, results in a thematic map showing the geographical distribution of a 'theme' such as vegetation or water quality. A classified remotely-sensed image is thus a form of a digital thematic map.

Spectral training is one of the most substantial barriers to traditional supervised parametric land cover classification. Analysts must laboriously develop a new set of training statistics for each new satellite image. The retraining is tedious, time consuming, expensive, and requires a trained image analyst.

Each information class of interest usually contains significant spectral variation, requiring several spectral subclasses for each class. This results in a large amount of processing time for a satellite image.

Research indicates that the common manual training techniques are inadequate, not only because of the analyst time requirements, but also because of incomplete information extraction from current generation satellite imagery (Bolstad and Lillesand, 1989). The accuracy of classification mainly depends on the quality of training rather than the applied statistical algorithm (Hixson et al, 1980). While a considerable effort has been devoted to the analysis of the performance of classification algorithms for use with remotely sensed multispectral images, relatively little attention has been focussed on the need to provide algorithms that would extract adequate and representative training data. The inclusion of non-representative pixels, in the training data, particularly in a heterogeneous environment, can seriously distort the sample statistics and hence the performance of the classifier.

Compared to imagery from first generation systems such as the Landsat MSS, current satellite have better spectral, spatial and radiometric resolution. For sensors with better spatial resolution, there will be many sub-classes in a single land use class, more overlapping between classes and thus less statistical separability between them. The tremendous spatial and spectral complexity of second generation data can make it extremely difficult and tedious both to locate and delineate spectrally homogeneous training fields. This is compounded by the need to ensure that all important subclasses within the high-variability data have been adequately characterized spectrally or even sampled at all (Arai, 1992). Because of these data induced changes, the application of traditional techniques to second generation satellite imagery requires a much greater time and expertise in order to produce accurate and repeatable classification results than has been typically the case with coarser, first generation data. And thus arises the need for evolving new and improved techniques for training area selection and purification. Therefore it is important that a representative training data be selected for feature

classification in the image. This particular aspect was the prime motivation for investigating some improved methods of training area definition and their application for the classification of digital remotely sensed imagery.

1.2 Thesis Objectives

The investigations in this thesis have been motivated by two main hypotheses:

Hypothesis I

The classification accuracy of high resolution scenes can be improved by seed pixel training area selection approach. This hypothesis underlines the first objective of the thesis.

Objective I

To evaluate the extent to which seed pixel algorithm for training area selection alleviate the classification problem.

Hypothesis II

Refinement of training area can improve the accuracy of classification. This hypothesis leads to the second and final objective of the thesis.

Objective II

To investigate some methods of training area refinement and to assess their significance in improving the classification accuracy.

1.3 Overview of Data Resources, Study Sites and Methodology

The study has been conducted on two study sites. The first is the Singrauli mines area in Sidhi district and the second is Bhopal city area, both in Madhya Pradesh (India). The four band Indian Remote Sensing Satellite (IRS) data have been obtained from National Remote Sensing Agency (NRSA), Hyderabad, India. The LISS2 IRS camera of the satellite operates in four band widths (0.45-0.52, 0.52-0.59, 0.62-0.68, 0.77-0.86 micrometers) with a spatial resolution 36.25m and a radiometric resolution of 7 bits per pixel. For defining training areas and testing the accuracies of final results, topographic maps (number 63 L/11, 63 L/12, 63 L/16, 1970 for Singrauli area and number 55 E/7 and 55 E/8, 1974 for Bhopal area) with a scale of 1:50000, published by the Survey of India, have been used.

The semi-natural environment of Singrauli area consists of five prominent classes. These are rocky terrain, dense forest, vegetation, water and mine dumps. The semi-urban environment of Bhopal city consists of five prominent classes. These are built-up area, open scrub/vegetation, marsh, forests and lake water.

The basic approach used in the thesis is to apply different training area extraction strategies, on study site images and to generate classification products. These products are

then compared to the test areas to compute classification accuracies. The investigation has involved writing computer programs. The list of main programs, which have been written in C language and with the help of Starbase Graphics libraries, have been given in Appendix.

1.4 Thesis Organization

Chapter 2 describes the Maximum Likelihood Classification and the training area extraction procedures with the traditional and improved methods. In Chapter 3, results of the experiments performed on imagery of Singrauli and Bhopal test area presented. Chapter 4 gives the concluding remarks with respect to the methods applied and the future perspective of training area refinement and extraction.

Chapter 2

TRAINING AREA EXTRACTION IN IMAGE CLASSIFICATION

2.1 Objective

In this chapter the theoretical background to per-pixel classification using supervised and unsupervised classification methods is presented with emphasis on supervised classification using maximum likelihood method. The choice of this algorithm is motivated by its proven performance for automatic classification of remotely-sensed images (Swain and Davis, 1978). The theoretical discussion is followed by a discussion on the various training area extraction methods and their limitations.

The objectives of the chapter are:

- To develop an understanding of the importance of 'good' training area selection in the per-pixel classification of the remotely-sensed imagery.
- To introduce an alternative approach for training area selection.
- To study some methods of training area refinement.

2.2 Supervised Image Classification

2.2.1 Overview

There are two types of image classification techniques supervised and unsupervised. In the supervised approach the user defines useful information categories and then examines their spectral separability whereas in the unsupervised approach he first determines spectrally separable classes and then defines their information utility.

In the supervised classification approach `the image analyst supervises the pixel categorization process by specifying to the computer algorithm, numerical descriptors of the various land cover types, called training areas or training sites, are used to compile a numerical interpretation key that describes the spectral attributes for each feature type of interest. Each pixel in the data set is then compared numerically to each category *it looks most alike*' (Lillesand and Keifer, 1987).

Generally, three algorithms are used for supervised classification: minimum-distance-to-means, parallelepiped, and Guassian maximum likelihood(ML) (Lillesand and Keifer, 1987). The minimum-distance-to-mean strategy is mathematically simple and computationally efficient but is insensitive to different degrees of variance or covariance in the spectral response of the training pixels. The parallelepiped classification strategy is also computationally simple and takes into account the variance in training classes but problems may arise from parallelepiped overlap due to correlation amongst classes. The Gaussian ML classifier quantitatively evaluates both the variance and covariance of the training class pixels and assumes a normal distribution for training classes. Of the three classifiers, the ML classifier has been most widely used. Hence, it has been selected in this study for supervised classification.

2.2.2 Maximum Likelihood Classification

Gaussian ML classifier quantitatively evaluates both the variance and covariance of training class pixels and assumes a normal distribution of training classes. This classifier is a parametric classifier which relies on specifying the probability distribution of class values by a representative function. In this case the function is Gaussian and the mean and covariance its parameters. The mean vector in multidimensional space locates the single peak of the unimodal probability distribution and the covariance specifies the region of influence of that probability distribution in each dimension. The classifier uses a parametric statistical covariance to prepare the probability density distribution function for each individual class and then employs Bayes' optimal strategy to maximize the likelihood of correct classification when allocating every pixel within the data set to one of the user specified classes. The user then decides upon an acceptable level of likelihood for each class and permits only those pixels above this level to be accepted into a final classification product. It is always assumed that prior to invoking the classifier the user has specified representative homogeneous training areas that permit an n-dimensional multivariate normal probability function to be constituted for each class. A more detailed discussion on the normal distribution is given by Hand (1981).

Since the normal distribution provides for good approximation for many naturally occurring phenomenon, a large amount of statistical theory has been based on it. It has been observed that a multivariate assumption holds well for the probabilistic processes observed in a large number of remote sensing applications. Also, classifiers designed on this basis are robust in the sense that classification is not very sensitive to even moderately severe violations of this assumption. From a practical point of view, experiments with more and less complex classifiers have shown that the normal assumption usually

provides a good trade-off between classification performance (accuracy) and cost (speed and complexity)(Swain and Davis, 1978).

In choosing the normal distribution for the ML classifier, two conditions must be satisfied. Firstly, the original data distribution should be unimodal and secondly, there should be an adequate number of training samples represented by a unimodal probability density distribution. The second major constraint is related to the need for an adequate number of training samples to be included in the classification. From the mathematical point of view, in an n-channel data set at least (n+1) samples should be provided for each class, otherwise the variance-covariance matrix for that class will become singular. In practice it is suggested that at least 10n to 100n pixels be sampled for each class (Swain and Davis, 1978).

Theoretically, the ML classification algorithm is known to be optimal in the sense of minimizing Bayesian error. However, in practice it has been shown that results produced by the classifier are not optimal and the classification results may show considerable variation depending upon the manner in which statistical classes are formed (Ince, 1987). The root cause for this lies in the underlying assumption of normality in the remotely-sensed data set. In many situations one may not even know how far off one may be from the true underlying distribution and therefore cannot estimate the consequences.

2.2.3 Mathematical Formulation for Software Operation

The probability P(x) that for a pixel vector x of p elements (a pattern defined in terms of p features) is a member of class ~~k~~ i is given by the multivariate normal density:

$$P(x) = 2\pi^{-0.5p} \left|\Sigma_i\right|^{-0.5} \exp[-0.5\,(y'\Sigma_i^{-1}y)] \qquad (2.1)$$

where |.| denotes the determinant of the specified matrix, Σ_i is the sample variance-covariance matrix for class i, $y = (x - \mu_i)$ and μ_i is the multivariate mean of class i. The term $(y'\Sigma_i^{-1}y)$ is the Mahalanobis distance, which measures the distance of an observation from the class mean, correlated for the variance and covariance of class i.

The function P(x) can be used to evaluate the probability that an unknown pattern x is a member of class i (i = 1,2,3...k). The maximum value in this set can be chosen and x allocated to the corresponding class. However, the cost of carrying out these computations for any but the most trivial case high. Savings can be made by first of all noting that we are only interested in the rank order of the values of P(x). Since the algorithm to the base e of a function has the same rank order as the function, the evaluation of the exponential term might be avoided by evaluating

$$\ln(P(x)) = -0.5p\ln(2\pi) - 0.5\ln\left|\Sigma_i\right| - 0.5(y'\Sigma_i^{-1}y) \qquad (2.2)$$

The rank order will also be unaffected if this expression is multiplied by 2 and the constant term pln2π dropped. The expression would look tidier if it was multiplied by -1 and the smallest value for all k classes chosen, rather than the largest. This would reduce the expression to

$$-2\ln(P(x)) = \ln\left|\Sigma_i\right| + y'\Sigma_i^{-1}y \qquad (2.3)$$

Further savings are made by computing the inverse and determinant of each Σ in advance and are read from a table when required. The computations then reduce to the derivation of Mahalanobis distance, the addition of the logarithm of the determinant of the estimated

variance-covariance matrix for each of the k classes in turn, and the selection of the minimum value from among the results. Since we have multiplied the original expression by -1 we minimize -2ln (P(x)) so as to achieve the same result as maximizing P(x). The computer time required increases exponentially with k, the number of classes, and with p, the number of bands(dimensionality) (Mather, 1987).

2.3 Role of Training Data in Supervised Classification

Supervised classification methods are based upon the prior knowledge of the number and the statistical parameters of the spectral classes. The statistical properties of the classes are estimated from the training pixels making up a class and the statistical parameters needed for classification will depend on the method used for the supervised classification. For example the parallelepiped method requires estimates of the extreme values on each of the feature for each class, while the k-means or centroid method needs estimates of the multivariate means of the classes. The method used in the thesis, the maximum likelihood algorithm, requires estimates of the mean vector and variance- covariance matrix of each class.

It is of crucial importance to ensure that the a priori knowledge of the number and statistical characteristics of the classes is reliable. The accuracy of a supervised classification analysis will depend upon two factors: (i) the representativeness of the estimates of the statistical spectral classes present in the image data and (ii) the degree of departure from the assumptions upon which the technique is based.

The validity of statistical estimates depends upon two factors - the size and representativeness of the sample. Sample size is related to the number of variables

(spectral bands in this case) whose statistical properties are to estimated, and the number of those statistical properties. For the multivariate property estimation the size should be atleast 30p pixels per class, where p is the number of features (spectral bands), and preferably more. The training samples are normally located by field work or from air photograph or map interpretation, and their positions on the image found by visual inspection.

2.4 Conventional Training Area Definition Procedures

The location of training areas in the image is normally established using windows, or portions of the full scene, in an enlarged format on an interactive color display device. The image analyst normally obtains training sample data by outlining training areas using a reference cursor. The cursor may be controlled by any of the several means (e.g. a tracker ball, joystick, digitizer, mouse, or keyboard stokes). The training areas are delineated on the image by drawing polygons on the image. The training area pixels are so located to avoid pixels along the edges between land cover types. The row and column coordinates of the vertices of these polygons are used as the basis for extracting (from the image file) the digital numbers of the pixel values located within each training area boundary. These pixel values then form the sample used to develop the statistical description of each training area. The conventional technique used for training area extraction has been shown in Figure 2.1. These training areas defined on the Bhopal and Singrauli images in band 3 are shown in Plates 2.1 and 2.2 respectively.

When delineating training set pixels, it is important to analyze several training sites throughout the scene. Training pixels for each class should be well distributed spatially to account for spatial variation in the class.

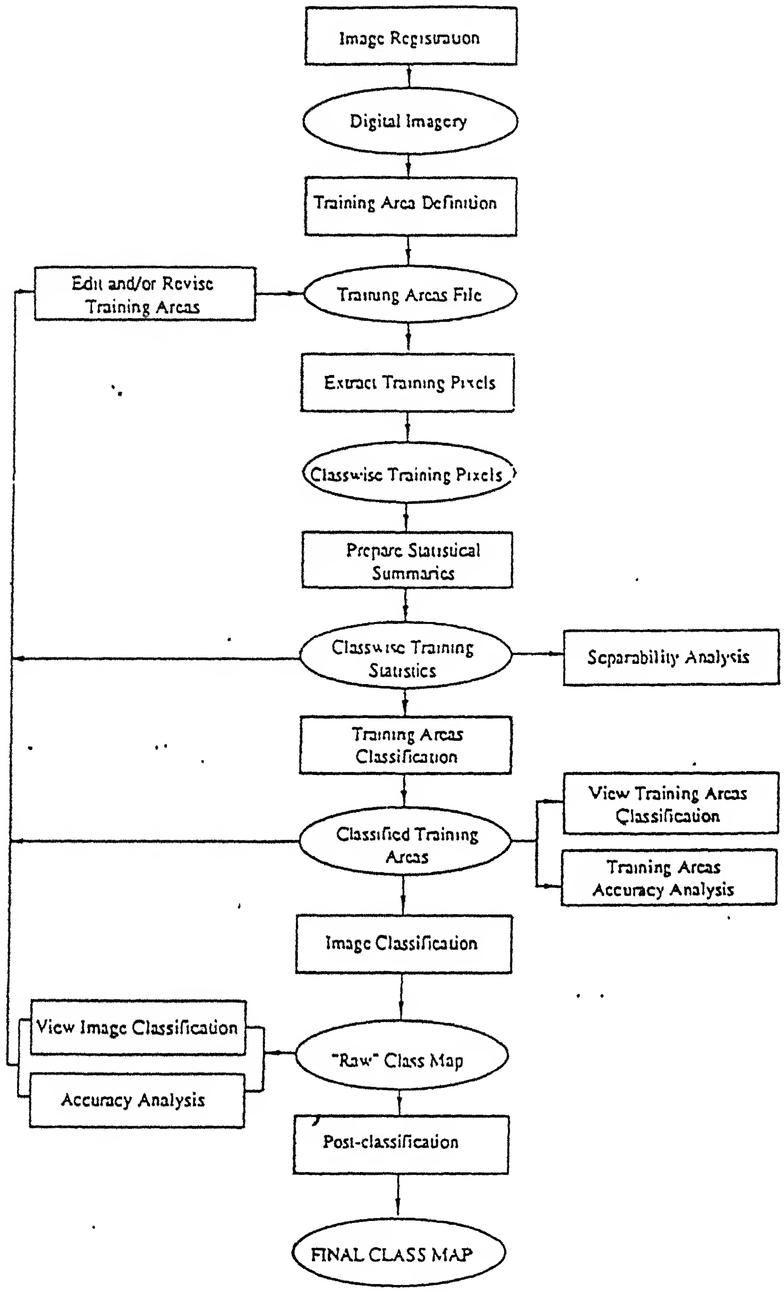

Figure 2.1: **Supervised Approach of Classification using the Gaussian ML Classifier (modified after Dikshit, 1992)**

Plate 2.1 Traditional training image superimposed on band 3 image, Bhopal Area.

Plate 2.2 Traditional training area image superimposed on band 3 image, Singrauli Area.

2.5 Evaluation of the Training Area

As part of the training area refinement process the overall quality of the training data contained in each of the original candidate training areas is assessed and the spectral separability between the data set is studied. One has to check if all data sets are essentially normally distributed and spectrally pure. Training areas that inadvertently include more than one spectral class are identified and recompiled. Likewise, extraneous pixels may be deleted from some of the data sets. These might be the edge pixels along class boundaries or within field pixels representing classes. Training sets that might be merged (or deleted) are identified, and the need to obtain additional training sets for poorly represented spectral class is addressed.

One or more of the following types of analysis is typically involved in the training set refinement process.

2.5.1 Divergence Analysis

A measure of the statistical separation between classes can be computed for all pairs of classes and can be presented in the from of a matrix. One statistical parameter commonly used for this purpose is divergence, a covariance-weighted distance between category means. In general, the larger the divergence, the greater the statistical separability and hence better accuracy in classification due to less confusion. Many mathematical expressions for divergence exist. One of the most widely used for calculating divergence is (Swain and Davis, 1978):

$$D_{i,j} = 0.5\ \mathrm{tr}(\Sigma_i - \Sigma_j) - (\Sigma_j^{-1} - \Sigma_i^{-1}) + 0.5\mathrm{tr}(\mu_i - \mu_j)^T - (\Sigma_i^{-1} - \Sigma_j^{-1})(\mu_i - \mu_j) \qquad (2.4)$$

where $D_{i,j}$ is the divergence between class i and j, and tr[A] is the trace of matrix A.

The range of D_{ij} is 0 to infinity, higher values implying greater separation. Because divergence increases without bound as statistical separability between classes reaches 100 percent, Swain et al. (1971) defined a saturation transformed divergence which provides a measure more closely corresponding to per cent classification. The pairwise transformed divergence is defined as

$$TD_{i,j} = K\,[1-\mathrm{Exp}(-D_{i,j}/8)] \tag{2.5}$$

K is a constant which can be taken as 100 to interpret transformed divergence as a percentage.

2.5.2 Classification of Training Data

Another evaluation of good training data is provided by a contingency table (which is also called a confusion matrix). This table is prepared by classifying the training set pixels. The known category types of the pixels used for training set are listed versus the categories chosen by the classifier. From this information, classification errors of omission and commission can be studied. In an ideal case, all nondiagonal elements of the contingency table would be zero, indicating no misclassification. Commission errors are represented by nondiagonal elements of the table where pixels are classified into a category to which they do actually belong. Omission errors represent the reverse type of situation. If more than an acceptable percentage of the pixels in a class is misclassified, that category may warrant further inspection and retraining.

The accuracy from contingency table based on training pixels should not be considered as a measure of overall classification accuracy. The contingency table simply tells us how well the classifier can classify the training areas and nothing more. Because the training areas are usually good, homogeneous examples of each cover type, they can be expected to be classified more accurately than less pure examples that may be found elsewhere in the scene. The overall accuracy can be evaluated only by considering test areas that are different from, and considerably more extensive than, the training areas.

The test for accuracy is usually in the form of a confusion table or error matrix or contingency table. This is an M x M table where M represents the total number of classes. The rows in the matrix represent the assumed true classes. The entries in the contingency table represent the raw number of pixels encountered in each condition and may also be expressed as percentages. One of the most important characteristic of such matrices is their ability to summarize errors of omission and commission. In the confusion table the diagonal elements represent the observations which agree on reference and classified training area and non-diagonal entries represent cases where observations do not agree.

2.6 Semi Automated Technique for Training Area Definition

2.6.1 Introduction

The traditional techniques for training area definition in supervised classification are manual techniques, i.e. the user selects an area on the image with the help of a mouse or any other similar device. The term 'semi automated' refers to training area selection automatically by the computer, with the user defining some conditions and parameters for

the region selection to the computer (and thus the term 'semi' gets introduced). The techniques used here were developed so as to increase the degree of automation in training field extraction, thereby significantly reducing analyst time requirement, while maintaining acceptable levels of classification accuracy. The procedure is still fundamentally grounded with 'per-point' (as opposed with 'per-field') classification framework, and therefore carries with it all problems inherent in such an approach (i.e., no spatial, textural, or contextual information is used in the classification). The intended purpose of the new technique, however, is to provide a tool whereby the utility of a per-pixel approach, limited as it may be for a second generation data applications, could be improved. It was also viewed as a potential complement to other non-traditional image analysis techniques and as a possible efficient conduit for integration of remotely sensed data and geographic information systems (GISs).

2.6.2 The Motivation to Evolve Alternate Methods for Training Area Selection

According to Swain and Davis (1978) the appropriate training sample size is between 10n and 100n, where n is the number of image channels. It can be quite easy to find homogeneous areas for crops when the field size is large, but the same task can be very difficult for natural vegetation, which usually exhibits great spatial and spectral variability.

The tremendous spatial and spectral complexity of second generation data can make it extremely difficult and tedious both to locate and delineate supervised training fields satisfying assumptions for the Guassian ML Classification. This is compounded by the need to ensure that all important - classes within the highly-variability data have been

adequately characterized spectrally, or even sampled at all. Furthermore, the clustering algorithms and parameters of many widely used unsupervised techniques were designed specifically with MSS data in mind, and thus are not universally applicable to data with fundamentally different characteristics. This factor is the major reason for the need to develop automatic procedures for extraction of training data. The extraction process should find those class-forming pixels, which are in majority inside a region, but have irregular ('spongy') spatial distribution. Semi-automated techniques may improve the efficiency and accuracy of spectral class identification. Region growing techniques have been developed which require the analyst to identify a single seed pixel for each training set (Buchheim, 1988). The algorithm then adds adjoining pixels to the training field in an option-controlled manner until one of several potential termination criteria is surpassed. The size, shape, spectral variability, and the local homogeneity can be used to control region growth. These techniques have been shown to improve training efficiency with no loss in classification accuracy over a range of conditions (Buchheim and Lillesand 1989).

2.6.3 An Alternative approach to Training Area Definition: The Seed Pixel Approach

As an alternative to manually delineating training area polygons, the semi-automated technique used here is the seed pixel approach to training. In this case, the display cursor is placed within a prospective training area and a single seed pixel is chosen so that it is thought to be representative of the surrounding area. Then according to various statistically based criteria, pixels with similar spectral characteristics that are four connected neighbours to the seed pixel are highlighted on the display and become the training samples of the training area.

In the seed growing technique for training area selection applied here, a 'seed' pixel is selected from a fairly homogeneous area (located visually), on the image. The algorithm then scans the four nearest neighbouring pixels. If the conditions

| Gray level of seed pixel - Gray level of the pixel being scanned | < User defined threshold

are satisfied, then that pixel becomes seed pixel, and it gets marked. Now, the four nearest neighbours of that pixel become seeds, and the above condition is again checked. This way, the recursive growth of pixels continues till all the seeds have all the four nearest neighbours not satisfying the thresholds. The order of evaluating the seeds is right, down, left and up, as shown in Figure 2.2

The seed algorithm works in the multispectral mode, i.e. it checks for the above growth condition in all the bands specified by the user, before marking a pixel as a training area pixel. Each seed pixel specified in the program may posses a unique threshold for its growth, such that, during training data extraction, the procedure can be sensitive to differences in within-class variation in spectral response that are often observed among cover types (e.g. water versus urban) in remotely sensed data. The training area selected by seed pixel algorithm on the two images is shown in Plate 2.3 and 2.4 respectively.

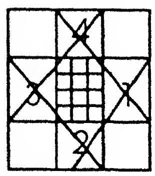

Figure 2.2 Order of growth of seed pixels in the algorithm used in the study. The centre pixel represents seed pixel defined by the user. The four nearest neighbours are marked by ⊠. The order of growth is indicated by numbers 1,2,3 and 4.

Plate 2.3 **Training area image generated by seed pixel approach, superimposed on Band 3 Image, Bhopal Area.**

Plate 2.4 **Training area image generated by seed pixel approach, superimposed on band 3 image, Singrauli Area.**

2.7 Training Data Refinement Techniques

2.7.1 Introduction

The fine line to be walked in the development of the training data sets is that of having sufficient sample size to ensure accurate determination of the statistical data parameters used by the classifier and to represent the total spectral variability in the scene, without going past a point of diminishing returns. In short, one does not want to omit any important spectral classes occurring in a scene, but also does not want to include redundant spectral classes in the classification process from a computational standpoint. During the process of training set refinement the analyst tries to identify such gaps and redundancies.

As part of the training set refinement process the overall quality of the data contained in each of the original candidate training areas is assessed and the spectral separability between the data set is studied. The analyst carefully checks if all data sets are essentially normally distributed and spectrally pure. Training areas that inadvertently include more than one spectral class are identified and recompiled. Likewise, extraneous pixels may be deleted from some of the data sets. These might be the edge pixels along class boundaries or within-field pixels containing classes or some other class pixels which might be present there. Training sets that might be merged (or deleted) are identified, and the need to obtain additional training sets for poorly represented spectral classes is addressed. The techniques used for the analysis of quality of the training area have been described above.

In general, the training set refinement process cannot be rushed with the 'maximum efficiency' attitude appropriate in the classification stage. It is normally an iterative procedure in which the analyst revises the statistical descriptions of the class types until

the classes are sufficiently separable. That is, the original set of candidate training area statistics is revised through merging deleting addition to form the final set of statistics used in the classification.

Training set refinement for the inexperienced data analyst may be a difficult task. Typically, an analyst has little difficulty in developing the statistics for the distinct nonoverlapping spectral classes present in a scene. If there are problems, they typically stem from spectral classes on the borders between the information classes - transition or overlapping classes. In such cases, the impact of alternative deletion and pooling of training classes can be tested by trial-and-error. In this process, the sample size, spectral variances, normality, and identity of the training sets should be rechecked. After studying the actual classification results, the image analyst might be faced with aggregation of certain detailed classes into more general ones.

One final note may be made here is that training set refinement is usually the key to improving the accuracy of a classification. However, if certain cover types occurring in the image have inherently similar patterns, no amount of retraining and refinement will make them spectrally separable! Alternative methods, perhaps including visual interpretation or fields checking, must be used to discriminate these cover types.

2.7.2 Necessity of Automatic Methods of Training Area Refinement

1. No enhancement procedure alone, can highlight the variability in the complete dynamic range of gray levels in the digital data. Some alien pixels are always bound to get included because of the limitation of the human eye to discriminate subtle color variations.

2. The training area set should be representative of the whole area to be classified and the derived parameters should fit to the applied classification model. The widely used maximum-likelihood (ML) classification requires the computation of mean value and covariance matrix for all categories. The sample should follow a Gaussian distribution for a reliable estimation of above parameters. Obtaining a Gaussian distribution of training data is facilitated by using homogeneous training areas. Homogeneity is usually judged by the user (interpreter) viewing the data in the form of a color composite on the screen of an image processing device. The visual check however can often be misleading because histogram equalization, which is a widely used enhancement method gives similar brightness for large values of pixel values lying at the margins of the actual dynamic range.

3. If a single band imagery were available, the study of a histogram of the frequency counts associated with each pixel value in the training data could help to reveal observations which were far from the mean (and hence the outliers). Alternatively, the cumulative frequency histograms might be plotted on normal probability paper, and outliers detected visually. Simple procedures such as these are not possible with high dimensional data.

4. Histogram can be used to display the distribution of the training area in each channel. When one of the distributions is multi-modal or long-tailed the user can modify the corresponding training area. However, this would increase the time-consuming process of training further; therefore an automatic cleaning of training data is recommend.

2.7.3 Automated Refinement Techniques

2.7.3.1 Overview

The inclusion of anomalous pixels in the training sites is a formidable problem in the supervised classification of images. Methods generally aim at overcoming these problems either by reconditioning the statistical parameters or by systematically dropping pixels from the training site data. The method proposed by Mather (1987) falls in the first category which aims at a robust estimation of statistical parameters by giving weights to pixel values in the training class depending on its Mahalanobis distance. In the second category a method proposed by Buttner et al. (1989) aims at iteratively removing pixels from the training data if their distance to the mean is greater than a certain predefined times standard deviation in any of the image bands.

Another method called visual outlier removal operates by eliminating pixels from the training site that lie outside a feature space mask. However, this method takes into account only two image bands at a time. The method proposed by Arai (1992) uses single band data to decide whether a particular pixel should be retained in that class or dropped. Gill and Gill (1993), presented a training site cleaning procedure which takes into account spectral overlap in all image bands and for all user specified classes simultaneously.

2.7.3.2 Mather's Method of Refinement

Importance of the method

- Use of this procedure is not limited to classification. Any statistical algorithm requiring estimates of mean and variance-covariance for a supposedly multivariate-normally distributed population would benefit from preprocessing using this method.

- The robust estimators are automatic, and the computing requirements are low.

- No visual interpretation of scatter plots is necessary.

- The procedure produces essentially the same estimates of μ and Σ and where 'pure' training data are supplied, and less biased estimates where the training data are contaminated.

Theoretical Foundation

The probability that an observed pixel vector x_0 belongs to class k given the sample estimates of μ and Σ for all possible classes is derived from the classical expression for the multivariate distribution:

$$p(x|w_k) = \frac{1}{2\pi^{p/2}\left|\Sigma_k\right|^{1/2}} \exp\left|1/2(x_0 - \mu k)' \Sigma_k^{-1} (x_o - \mu k)\right| \tag{2.6}$$

where μ_k and Σ_k are the estimated mean vector and variance-covariance matrix for class k, p is the number of spectral bands, and Σ_k^{-1} is the inverse of Σ_k. The expression $p(x_0|w_k)$ are calculated for all m classes; pixel vector x_0 is allocated to that class for which $p(x_0|w_k)$ is a maximum.

The parameters of the above are μ_k and Σ_k. The former fixes the centre (location) of the probability distribution of associated with each class k, while the latter defines the shape of distribution. If either estimate is inaccurate then the reliability of the calculated probabilities $p(x_0|w_k)$ will be affected. One major cause of inaccuracy in estimating μ and Σ is the inclusion of pixel vectors in the training data for class k which, in reality, do not belong to class k; they may belong to another class altogether or they may be hybrid or mixed pixels.

Mather's technique is based on a paper by Campbell (1980), which has considered ways in which atypical values in a sample could be detected and proposes estimators of mean and variance-covariance matrix which are robust (that is, they are not unduly influenced by atypical values). These estimators give full weight to observations that are assumed to come from the main body of the data but reduce the weight given to observations identified as aberrant. A measure called Mahalanobis distance is used to identify deviant members of the sample. Its square is defined by

$$d_m^2 = (x_i - \mu_i)' \Sigma^{-1} (x_i - \mu_i) \qquad (2.7)$$

where i is the index counting the elements of the sample, x_i is the i^{th} sample value (pixel vector), μ_i is the mean vector and Σ is the sample variance-covariance matrix of a given sample (class). The transpose of vector x is denoted by x'. Robust estimates of mean and variance-covariance matrix are computed using weights which are functions of

Mahalanobis distance. The effect is to downgrade those pixel values with high Mahalanobis distances, which represent pixels that are relatively far from (dissimilar to) the mean of training class, taking into account the shape of probability distribution of training class members. For uncontaminated data the robust estimates are close to those obtained from the usual estimators. The procedure for obtaining the weights is described and illustrated by Campbell (1980);

$$\mu_k = \frac{\sum_{i=1}^{n} w_i x_{ik}}{\sum_{i=1}^{n} w_i} \qquad (k = 1,2,......,p) \tag{2.8}$$

$$\Sigma_{ik} = \sum_{i=1}^{n} w_k^2 (x_{ki} - \mu_i)(x_{kj} - \mu_j) / \left(\sum_{i=1}^{n} w_k^2 - 1 \right) \tag{2.9}$$

where

n = number of pixels in the training sample

p = number of features

w_i = weight per pixel

x_{ki} = value for pixel on feature k

x_k = mean of k th feature for this case

s_{jk} = j, k th element of variance-covariance matrix for this class

The weights are found from:

$$w_i = F(d_i)/d_i \tag{2.10}$$

given

$$F(d_i) = \begin{cases} d_i & \text{if } d_i \leq d_0 \\ d_0 \exp[-0.5(d_i - d_0)^2 / b_2^2] & \text{otherwise} \end{cases} \quad (2.11)$$

and

d_i = Mahalanobis distance of pixel i for this class

$d_0 = p + b_1/2$

$b_1 = 2$

$b_2 = 1.25$

The weights w_i are initially computed from the Mahalalnobis distances which are themselves computed from μj and Σ_j derived from the above formulae with unit weights. The Mahalanobis distance and the unit weights are recalculated iteratively until successive weight vectors converge to within acceptable limits. On the basis of experiments, Mather (1987) defined the convergence as being reached when the maximum absolute change in any weight w_i from one iteration to the next was less than 0.01. It is then found that any aberrant pixel vectors is given very low weights and will therefore contribute only negligibly to the final (robust) estimates of μ and Σ which are required in the maximum likelihood classification scheme.

2.7.3.2 Buttner's Method of Refinement

Important Features

It actually cleans the data, i.e. it removes the non-representative pixels from the training data. It checks for the pixel removal condition in all bands, i.e. if a pixel is representative in one band but an outlier in any other band it is removed from the training data. The algorithm is simple and fast. The cleaning can be achieved in few minutes for digital imagery in three bands.

Background

It is a method to improve the effectiveness of the supervised learning in an interactive environment. An iterative procedure for cleaning training data is introduced to provide more appropriate class statistics (in terms of normality criterion) for classification. The anomalous pixels in the class are removed in a simple automatic way.

Pixels are removed iteratively from the training data if their distance to the actual mean value is greater than a predefined limit. Considering the case of ML Classification this corresponds to a low probability of belonging to the class defined by μ and Σ. All pixels are taken into account in each iteration step.

In the actual implementation the pixel $x(x_1, x_2,..,.x_i,...,x_n)$ is omitted if one of the following inequalities is fulfilled

$$|x_i - M_i| > k\sigma_i \qquad i = 1,2, ...n \qquad (2.12)$$

where M_i is the i^{th} element of the estimated mean vector, σ_i is the ith element of the estimated standard deviation, x_i is the ith element of the pixel vector, n is the number of channels and k is a user-specified parameter for the limit. The iteration is stopped if the maximum change in standard deviations is less than a limit

$$(\max |(\sigma_i)^j - (\sigma_i)^{j-1}|) < 0.1 \qquad (2.13)$$

where j is the iteration number. A maximum of ten iterations are allowed. By removing the outlying pixels, the distribution in each channel will approximate better to Gaussian, which in turn results in a better parameter estimation for the multi-dimensional normal distribution. The pixels removed by this method when refinement is applied on a conventional traininig area (selected as a box) has been shown in Plate 2.5.

The limitation of this method is that it cleans a training site without taking into consideration the spectral overlap with other classes.

Plate 2.5 **The pixels removed by Buttner's technique when applied on a conventional training area.**

Chapter 3

RESULTS AND DISCUSSION

3.1 Introduction

In chapter 2, the importance and methodology of alternate training area extraction procedure were discussed. The results of these methods have been presented in this chapter. The experiments have been performed on imagery of two sites, one of Singrauli mine area and the other of Bhopal city area. The classification of these sites using the traditional and improved methods have been discussed in this chapter. The classes used in classification have been listed in Table 3.1.

3.2 Experimental Methodology

The training area extraction and analysis was done under the following steps:

- The training areas for classification of imagery was selected by conventional manual technique of drawing boxes on the image. The ML classification was carried out and the individual class accuracies were established.

- In the second step, the training areas for classification of imagery were selected by the seed pixel algorithm, with the initial seeds in those areas where the traditional training areas had been selected. In the seed pixel algorithm used here, the user provides a threshold gray level to be included in the region, based on the mean and standard deviation of the pixels in a ~~3~~ 7 x ~~3~~ 7 window. The ML classification was applied to this training area also.

- The training areas selected by the above two techniques were subjected to refinement techniques by two methods: the Buttner's approach of removing aberrant pixels from the class, at different standard deviations factors (k=1,2,..) as the threshold limit; and the Mather's method (Mather, 1987) of calculating robust statistics for classification. These two methods have been discussed in the section 2.8.2 and 2.8.3 respectively in the second chapter. Classification was performed on the original training data by the Mather's method and also on the refined data obtained by Buttner's approach of removing the aberrant pixels from the original data. The accuracy of classification of training data and test data was studied in each case.

An overview of the methodology has been shown below in the Figure 3.1.

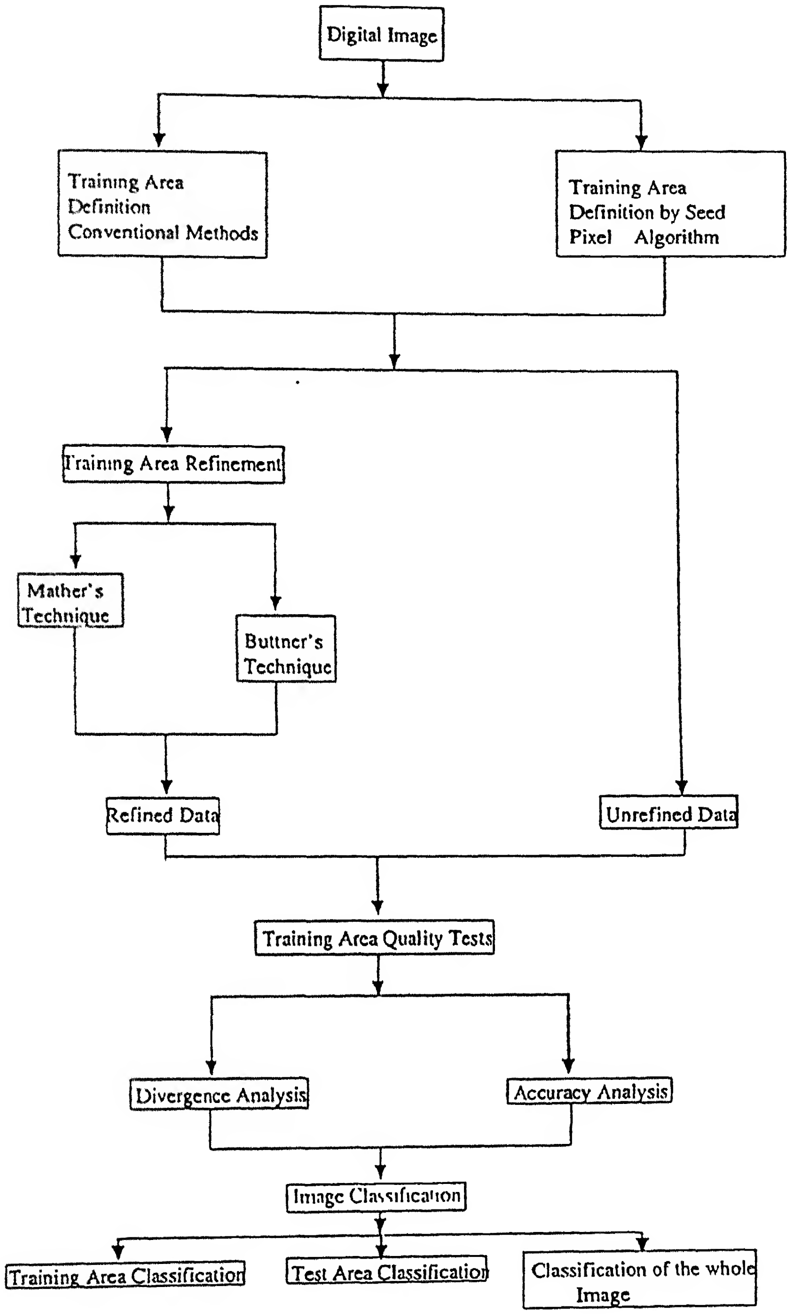

Figure 3.1 Flowchart showing the overall methodology followed in the study.

3.3 RESULTS

3.3.1 Experiments with Bhopal City Imagery

The LISS II image of Bhopal city was used as the first test image. The semi-urban environment of Bhopal city consists of five prominent classes. Water, the main class in the image is mainly represented by the two big lakes in the city. The lakes are surrounded by marshy land. The residential area of the city is flanked by forests and unused land marked by sparse vegetation and scrub.

The classification results can be analyzed by considering the following factors:

a. Number of Pixels

The number of pixels being selected as a training area can be controlled by the analyst in the traditional techniques, but in the seed pixel approach, the number of pixels getting marked as a training area is controlled by the algorithm being used for seed growth. The number of training area pixels selected by traditional techniques is 611 in the traditional case, while 4534 in the seed pixel approach.

b. Mean and Standard Deviation

The mean and standard deviation of the classes after each refinement step have been given in Table 3.2. It is observed that standard deviation in all the image bands for all the classes reduces after both the methods of refinement. This is associated with a shift in class means, which is likely to improve inter-class separability.

c. Class Covariance

The class covariance for each class is presented in Table 3.4. It is observed that the covariance values from the traditional techniques are inflated by contamination. Therefore, the multivariate-normal probabilities calculated using initial estimates of Σ would be in error. Consequently the performance of the ML classifier will be reduced. It was observed that the estimates of Σ after the Mather and Buttner refinement is low, indicating the lesser spread of the class, and less aberrant pixels in it.

d. Divergence Value

Average divergence value increases after refinement method proposed by Buttner, indicating the greater separation between classes. The average, minimum, and maximum divergence values of traditional training area were 559, 35 and 1960 respectively. After refinement by the Buttner's method at k=2, the average divergence value increased to 630, with minimum and maximum values became 41 and 2101; This shows an increase in the separability after refinement between classes. The minimum transformed divergence increased to 99 from 98, while the maximum remained 100 before and after refinement. The average transformed divergence values did not show change, indicating that a sufficient separability had been achieved for classification before refinement (Tables 3.6 and 3.7) .

e. Accuracy of Classification

The following observations were made:

1. Classification accuracies for unrefined training area using seed pixel approach is higher than that of the traditional technique.

2. The Buttner's method refinement at k=2 proved to be the best for ML classification in this case. While k=1, removes the inherent heterogeneity of a class, k=3 value tends to include aberrant pixels which diminish the accuracy of classification (Table 3.13).

The individual class accuracies have been analyzed below:

Water This class is most accurately classified by all the methods applied in the study. The test and training area classification accuracy is more than 99% in nearly all the cases.

Forest The traditional training area technique gives the test accuracy of 80.37%. This accuracy falls sharply by using the Buttner's technique at k=1, when many pixels have been dropped after the refinement process. At k=2, when most of the pixels in the class have been retained, the accuracy improves, but remains low compared to the traditional technique. It may be that in a heterogeneous class like forest, the statistically aberrant pixels are an important member of the class. These pixels should also be included in defining/capturing the overall statistical variability of a heterogeneous class. The Mather's refinement procedure gives the same accuracy as that of the traditional technique.

The seed pixel algorithm sharply improves the test area classification accuracy of the classifier (97.65%). It shows that seed algorithm can be very useful in heterogeneous environment, where it can select the classes in a much better way than the traditional methods.

Marsh It is the most poorly classified class by all the strategies applied here. The traditional training area selection techniques gives the test area classification accuracy of 55.12% . The Buttner refinement technique at k=1 was able to improve it to 61.33%, which was the highest test accuracy attained by any of the methods used here for this class. This bad classification may be due to the fact that marsh is a transition class, representing a land - water boundary.

Urban Area The Buttner's refinement technique at k=2, was able to improve the test classification accuracy by traditional methods (89%) to 91.5%.

Scrub The Buttner's refinement at k=2, improved the classification accuracy by traditional technique (89%) to 91.5%.

3.3.2 Experiments with Imagery of Singrauli Area

The LISS II image of Singrauli mine represents a semi-natural environment. The image consists of rocky terrain and forests. There are large mine dump areas on the image. In the south-eastern part of the image, the Govind Ballabh Pant reservoir is present. There are many water filled areas on the image too. Since the banks of the reservoir is highly silted, it has been considered as a different class from clear water bodies.

a. Number of Pixels

The number of pixels selected for training by seed algorithm is quite high (7015) compared to the training area pixels selected by conventional techniques (1456). The main reason is the selection of reservoir water class training area by the seed algorithm. Since

the standard deviation of the reservoir water class is quite low, the seed algorithm selects nearly the whole reservoir water area as a training site.

b. Mean and Standard Deviation

As expected, the mean and standard deviation decrease with each refinement step by the Buttner's refinement technique. A step by step change in standard deviation at different k values is given in Table 3.3.

c. Covariance

The covariance values before and after the refinement by the Buttner's and the Mather's approach have been given in Table 3.5. As expected, the refined covariance values are lower than the original ones.

d. Divergence

It was found that the seed algorithm improved the inter-class separability for each individual class (Tables 3.10 and 3.11), although the average divergence remained lower than that obtained by traditional method. The average divergence values increased after refinement by Buttner's technique, but the transformed divergence remained constant at 99 due to its saturating behaviour.

e. Accuracy of Classification

The following observations were made:

1. The Mather's technique of refinement does not show any significant change in classification accuracy when compared to the traditional classifier.

2. Application of Buttner refinement technique on the seed pixel training area actually decreased the classification accuracy rather than increasing it at some k values. Mather's technique showed marginal improvement in accuracy when applied on the seed training data.

The individual class accuracies have been analyzed below:

Reservoir Water This class is having 100% accuracy of classification for training area as well as test area by all the classification strategies applied here.

Forest Training Area classification accuracy by traditional technique is 94.55%. With the seed pixel algorithm approach it increased to 100%. Test area accuracy is maximum (86.21) by the Buttner refinement of traditional data at k=2. The seed pixel algorithm gives a comparatively low classification accuracy (70.88%).

Mine Dumps The training area accuracy is 100% by all the methods applied. The traditional technique gave the test area classification accuracy as 100% before and after the Buttner's refinement technique whereas the seed pixel algorithm gave lower classification accuracy of 97.74% for the same case.

Rocky Terrain The training area classification accuracy of 100% was obtained by all methods. But the maximum test area accuracy achieved by traditional methods is 98.62% , which improves marginally to 98.90% by Mather's method. The maximum accuracy

achieved in this case is by seed pixel algorithm, which gives 100% test area accuracy for this class.

Water Bodies Training area accuracies obtained for this class is 100% by all the methods, except when the refinement by Buttner's technique is applied at k=1. Maximum test area accuracy is achieved by seed algorithm (96.94%). While the traditional technique gave classification accuracy of 84.62%, the Mather's technique marginally improved it to 84.97%.

3.4 DISCUSSION

For better training area extraction, alternate training area selection procedures and refinement techniques were used in the thesis. Although, the overall classification accuracy did not show sharp improvement by any of the methods used, but the individual class accuracies, showed marked changes with refinement.

3.4.1 Refinement using Robust Estimates of Training Statistics

In this refinement technique proposed by Mather (1987), the robust estimators are automatic, and computing requirements are low (the ML classification took nearly 4 minutes for classification for 5 classes, while the refinement and classification by Mather's technique took 5 minutes for the same case, on HP 9000/735 systems). No visual interpretation of scatter plots is necessary. The procedure produces essentially the same estimates of μ and Σ where *pure* training data are supplied, and robust estimates where the training data are contaminated. The efficiency of the proposed procedure is

demonstrated by the fact that the classification accuracy by this approach matches with that of the ML classifier, and marginally improves the ML classification results when the classifier accuracy is already high. This method takes into account the variance-covariance in all image bands.

In this technique the weights are a function of D_i, the Mahalanobis distance. which in turn is a function of D_0, i.e. $D_i = \alpha D_0$ (equation 2.11). Mather suggested the value of α to be 0.75. He derived this value by experimenting with many synthetic images. This factor was found to be suitable for natural images used in this investigation also. It was observed that the accuracy of classification falls sharply with the change of α. The change of D_0 factor from 0.75 to 0.70 reduced the accuracy of training area classification from 98% to 66% for the Singrauli mine image.

This method took a maximum of 6 iterations for upgrading the training area statistics in experiments. This less number may be due to the fact that the original data were already having high separability.

3.4.2 Refinement by Aberrant Pixel Removal

This simple method of improving the effectiveness of supervised training in an interactive environment, proposed by Buttner et. al. (1989) gave significant changes in standard deviation, divergence and accuracy of classes. The following observations can be made:

1. Less than 10 iterations were enough to stop the refinement process in each class.

2. The lower k value required a larger number of iterations.

3. Changes in mean and standard deviations are more in infrared channels than in visible ones.

4. The differences are very small in the mean and standard deviation values for water due to small within class variance, while they are pronounced for vegetation categories (larger within-class variance). Moderate changes in class stability are obtained for rocky terrain class.

5. In the vegetation classes, the classification accuracy drops at k=1, and improves sharply at k=2. Because of large standard deviation values only a few pixels are considered to be outlying at k=2.

6. The benefit of iterative training is seen from the improvement in class separability. The improvement is mainly due to the change of the standard deviation, since as shown in Tables 3.6, 3.7, 3.8, 3.9, 3.10, 3.11, 3.12, the mean values are rather insensitive to the chosen values of the parameter k. Therefore, classification methods using second order statistics, such as the ML algorithm, can benefit from the iterative training.

3.4.3 Seed Pixel Algorithm

By virtue of its design and implementation, the seed pixel procedure of training area definition has the potential to solve several of the problems encountered with traditional forms of supervised classification with second generation satellite imagery:

1. Because only one pixel must be located in the image, the seed pixel algorithm reduces the probability of wrongly locating the ground classes on the image and may produce dramatic reductions in the amount of analyst time required for training field delineation. Thus this algorithm, may eliminate much of the complexity involved in training on spectrally and spatially complex imagery.

2. In contrast to the traditional supervised procedures which mark pixels using rectangular windows of image data, pixel extraction using seed algorithm strategy is not biased toward the exclusion of objects of a particular shape or orientation: fields may grow out along linear features as easily as within rectangular features; the importance of adequately characterizing such features spectrally has increased with the improved spatial resolution of second generation data.

3. In the seed pixel algorithm, each seed pixel specified in the program may possess a unique threshold for its growth, such that, during training data extraction, the procedure can be sensitive to differences in within-class variation in spectral response that are often observed among cover types (e.g. water versus urban) in remotely sensed data.

4. The seed pixel algorithm may be used very effectively to delineate the large numbers of training fields which may be required to adequately characterize the spectral variability of second generation satellite data; manual marking of such data can be exceedingly tedious.

The results achieved for overall classification accuracy by the seed pixel algorithm used in this work show marked improvement in training area accuracy (100% accuracy attained in almost all cases). Although the test area classification accuracy is quite high (more than 90%), it is marginally better in its performance when compared to the traditional technique results. When individual class accuracies are analyzed, the seed pixel algorithm gives better results in nearly all cases, but the marked improvement in its performance is only shown in classifying natural heterogeneous classes like forests and rocky terrain. The classification accuracy with seed pixel algorithm in the test areas for these classes increases to 97.65% and 100% respectively as compared to the traditional technique results of 80.37% and 98.62% respectively.

Table 3.1 **Classes used in classifying the images of the two study sites.**

Bhopal Image

Class No.	Information Class
1.	water
2.	forest
3.	marsh
4.	urban (built-up area)
5.	scrub

Singrauli Image

Class No.	Information Class
1.	reservoir water
2.	forest
3.	mine dumps
4.	rocky terrain
5.	water bodies

Table 3.2 **Mean and standard deviations obtained with different k parameter values for the refinement applied by the Buttner's technique on the traditional training area of Bhopal Image**

Class	k	Band 3		Band 4		Band 1		Iteration #
		μ	σ	μ	σ	μ	σ	
Water	∞	16.0	0.0	11.45	0.00	34.64	0.03	
	2	16.0	0.0	11.45	0.00	34.64	0.03	1
	1	16.0	0.0	11.45	0.0	34.64	0.03	1
Forest	∞	23.65	1.832	27.70	2.256	37.42	0.96	
	2	23.42	1.40	27.96	0.0	32.33	0.81	1
	1	24.0	0.0	28.0	0.0	37.0	0.0	3
Marsh	∞	18.23	0.639	22.839	1.713	34.167	0.628	
	2	18.23	0.0	22.839	1.713	34.167	0.628	1
	1	23.06	0.807	34.0	0.0	18.0	0.0	2
Urban	∞	31.295	1.167	29.261	1.419	42.88	1.19	
	2	31.18	1.003	29.17	1.305	42.85	1.02	1
	1	31.0	0.0	29.0	0.0	43.0	0.0	2
Scrub	∞	38.673	1.660	41.465	1.158	42.54	1.05	
	2	38.85	1.49	41.65	0.976	42.5	0.97	1
	1	39.09	0.792	41.56	0.496	42.62	0.485	3

Table 3.3 Mean and standard deviations obtained with different k parameter values for the refinement applied by the Buttner's technique on the traditional training area of Singrauli Image

Class	k	Band 1		Band 3		Band 4		Iteration Number
		μ	σ	μ	σ	μ	σ	
RW*	∞	86.74	0.605	79.80	0.529	34.77	0.593	
	2.5	86.74	0.605	79.80	0.529	34.77	0.593	1
	1	87.00	0.000	80.00	0.000	35.00	0.00	1
Forest	∞	64.13	2.600	46.40	3.41	44.123	6.80	
	2.5	64.13	2.600	46.239	3.22	44.123	6.792	1
	1	63.00	0.000	46.00	0.00	43.4	0.49	4
Mine	∞	62.37	1.755	52.575	2.82	58.9	4.34	
	2.5	62.37	1.755	52.556	2.799	59.9	4.34	1
	1	63.63	0.481	53.00	0.00	59.49	0.50	3
Rock	∞	105.61	10.899	109.79	11.27	92.37	13.22	
	2.5	105.61	10.899	111.59	8.35	92.373	13.225	1
	1	103.00	0.000	113.00	0.00	98.00	0.00	6
Water	∞	74.496	3.017	80.97	5.04	75.62	3.68	
	2.5	74.496	3.017	81.127	4.852	75.908	3.292	1
	1	75.000	0.000	82.00	0.00	76.00	0.00	5

*RW : Reservoir Water

Table 3.4 **The change in covariance values with each iteration for each class in the Mather's refinement technique applied on the Bhopal Image. The decrease in the covariance values with increasing iteration is observed.**

Class No.	Iteration No.	Covariance Matrix
1	1	1.526 0.180 0.987 0.180 0.311 0.138 0.987 0.138 1.062
	2	1.480 0.173 0.946 0.173 0.311 0.133 0.946 0.133 1.028
2	1	4.729 -12.698 1.685 -12.698 52.403 -5.210 1.685 -5.210 0.999
	2	4.718 -12.812 1.681 -12.812 52.182 -5.280 1.681 -5.280 0.996
	3	4.707 -12.823 1.678 -12.823 52.170 -5.288 1.678 -5.288 0.995
3	1	1.067 1.992 0.369 1.992 9.815 1.123 0.369 1.123 0.513
	2	1.061 1.962 0.344 1.962 9.733 1.039 0.344 1.039 0.459
	3	1.060 1.956 0.340 1.956 9.716 1.022 0.340 1.022 0.448
4	1	3.047 -3.071 2.553 -3.071 34.652 -5.637 2.553 -5.637 2.976
	2	2.982 -2.932 2.520 -2.932 34.474 -5.573 2.520 -5.573 2.963
	3	2.966 -2.897 2.512 -2.897 34.429 -5.558 2.512 -5.558 2.960
5	1	11.043 7.000 -2.473 7.000 6.237 -1.236 -2.473 -1.236 1.512

Table 3.5 The change in covariance values with each iteration for each class in the Mather's refinement technique applied on the Singrauli Image.

Class No.	Iteration No.	Covariance Matrix		
1	1	0.280	0.076	0.081
		0.076	0.352	0.156
		0.081	0.156	0.367
	2	0.279	0.080	0.083
		0.080	0.343	0.150
		0.083	0.150	0.364
	3	0.279	0.080	0.083
		0.080	0.342	0.150
		0.083	0.150	0.364
2	1	11.758	14.638	6.080
		14.638	46.419	-0.215
		6.080	-0.215	6.863
	2	11.643	14.572	5.937
		14.572	46.364	-0.382
		5.937	-0.382	6.711
	3	11.611	14.556	5.900
		14.556	46.359	-0.417
		5.900	-0.417	6.673
3	1	7.970	6.438	3.971
		6.438	18.916	2.671
		3.971	2.671	3.088
	2	7.873	6.681	3.934
		6.681	18.585	2.775
		3.934	2.775	3.077
	3	7.860	6.738	3.927
		6.738	18.480	2.806
		3.927	2.806	3.074
4	1	128.129	89.674	91.817
		89.674	176.389	7.308
		91.817	7.308	119.812
	2	126.103	88.254	90.627
		88.254	174.554	4.129
		90.627	4.129	118.227
	3	125.970	88.111	90.506
		88.111	174.075	2.931
		90.506	2.931	117.820
	4	125.913	87.934	90.344
		87.934	173.709	2.135
		90.344	2.135	117.478
	5	125.885	87.822	90.241
		87.822	173.490	1.676
		90.241	1.676	117.272
5	1	25.518	14.235	11.281
		14.235	13.586	7.662
		11.281	7.662	9.124
	2	25.366	14.042	11.251
		14.042	13.337	7.614
		11.251	7.614	9.127

Table 3.10 Training area statistics and divergence analysis of the traditional training area for the Singrauli Image

Image Bands Used are 1,3,4.

Mean Values

Class	No. of Pixels	1	2	3
1	400	86.74	79.80	34.77
2	162	64.14	46.41	44.12
3	409	62.38	52.57	58.90
4	118	105.61	109.80	92.37
5	367	74.50	80.97	75.62

Class Covariance

1	.37	.08	.16
	.08	.28	.08
	.16	.08	.35
2	6.86	6.08	-.22
	6.08	11.76	14.64
	-.22	14.64	46.42
3	3.09	3.97	2.67
	3.97	7.97	6.44
	2.67	6.44	18.92
4	119.81	91.82	7.31
	91.82	128.13	89.67
	7.31	89.67	176.39
5	9.12	11.28	7.66
	11.28	25.52	14.24
	7.66	14.24	13.59

Divergence Matrix

	CLASSES				
CLASSES	1	2	3	4	5
1	0	3336	4457	5457	4131
2	3336	0	20	203	134
3	4457	20	0	317	70
4	5457	203	317	0	64
5	4131	134	70	64	0

The Average Divergence Value = 1819

Transormed Divergence Matrx

	CLASSES				
CLASSES	1	2	3	4	5
1	0	100	100	100	100
2	100	0	92	100	99
3	100	92	0	100	99
4	100	100	100	0	99
5	100	99	99	99	0

The Average Transformed Divergence Value = 99

Table 3.11 Training area statistics and Divergence Analysis of the Singrauli Image at k=1.

Image Bands Used are 1,3,4.

Mean Values

Class	No. of Pixels	1	2	3
1	128	86.62	79.38	34.70
2	143	64.30	46.46	43.68
3	352	62.26	52.51	58.77
4	111	105.77	109.59	91.95
5	333	74.51	80.87	75.62

Class Covariance

1	.46	.18	.25
	.18	.61	.19
	.25	.19	.43
2	7.21	6.82	1.15
	6.82	13.31	16.80
	1.15	16.80	47.95
3	3.31	4.56	3.16
	4.56	9.23	7.42
	3.16	7.42	21.07
4	126.89	98.23	8.96
	98.23	135.59	93.91
	8.96	93.91	184.43
5	9.61	12.46	8.26
	12.46	28.01	15.69
	8.26	15.69	14.45

Divergence Matrix

CLASSES	1	2	3	4	5
1	0	2374	4044	4165	4291
2	2374	0	21	183	125
3	4044	21	0	300	61
4	4165	183	300	0	64
5	4291	125	61	64	0

The Average Divergence Value = 1563

Transormed Divergence Matrx

CLASSES	1	2	3	4	5
1	0	100	100	100	100
2	100	0	92	100	99
3	100	92	0	100	99
4	100	100	100	0	99
5	100	99	99	99	0

The Average Transformed Divergence Value = 99

Table 3.12 Seed pixel training area statistics and divergence analysis of the for the Singrauli Image

Image Bands Used are 1,3,4.

Mean Values

Class	No of Pixels	1	2	3
1	5044	86.27	80.80	37.70
2	212	63.28	53.09	60.83
3	283	105.75	115.78	99.94
4	1226	73.87	79.29	77.64
5	250	63.20	48.16	51.25

Class Covariance

1	3.38	.84	-2.26
	.84	2.05	1.28
	-2.26	1.28	8.43
2	.65	.40	-.23
	.40	1.30	-.10
	-.23	-.10	1.78
3	17.25	11.91	6.99
	11.91	17.32	10.91
	6.99	10.91	11.65
4	7.99	1.83	1.11
	1.83	6.37	1.42
	1.11	1.42	2.88
5	3.03	2.23	-1.33
	2.23	4.20	-.38
	-1.33	-.38	8.91

Divergence Matrix

	CLASSES				
CLASSES	1	2	3	4	5
1	0	864	707	465	475
2	864	0	2958	462	46
3	707	2958	0	205	904
4	465	462	205	0	315
5	475	46	904	315	0

The Average Divergence Value = 740

Transormed Divergence Matrx

	CLASSES				
CLASSES	1	2	3	4	5
1	0	100	100	100	100
2	100	0	100	100	99
3	100	100	0	100	100
4	100	100	100	0	100
5	100	99	100	100	0

The Average Transformed Divergence Value = 99

	Class	Pixels	Cor_Pix	% Corr		Class	Pixels	Cor_Pix	% Corr			Class
	1	128	128	100.00		1	400	400	100.00			1
	2	143	64	44.76		2	159	150	94.34			2
	3	352	352	100.00		3	408	404	99.02			3
	4	111	111	100.00		4	110	110	100.00			4
	5	333	180	54.05		5	363	363	100.00			5
	Total	1067	835	78.26		Total	1440	1427	99.10			Total
TRAD	k=1					k=3					SEED	k=1
	Class	Pixels	Cor_Pix	% Corr		Class	Pixels	Cor_Pix	% Corr			Class
	1	282	282	100.00		1	282		100.00			1
	2	371	316	85.17		2	371		86.25			2
	3	311	311	100.00		3	311		100.00			3
	4	364	208	57.14		4	364		98.89			4
	5	426	245	57.51		5	426		85.25			5
TEST	Total		1362	65.05		Total			93.21			Total

Pixels	Cor_Pix	% Corr		Class	Pixels	Cor_Pix	% Corr		Class	Pixels	Cor_Pix	% Corr
3421	3421	100.00		1	4897	4897	100.00		1	400	400	100.00
138	138	100.00		2	212	212	100.00		2	162	153	94.44
257	257	100.00		3	283	283	100.00		3	409	405	99.02
1077	1077	100.00		4	1226	1226	100.00		4	118	118	100.00
203	191	94.09		5	249	249	100.00		5	367	367	100.00
5096	5084	99.77		Total	6867	6867	100.00		Total	1456	1443	99.11
				k=3					UNREF	TRAD		
Pixels	Cor_Pix	% Corr		Class	Pixels	Cor_Pix	% Corr		Class	Pixels	Cor_Pix	
282	282	100.00		1	282	282	100.00		1	282	282	100.00
371	305	82.21		2	371	259	69.81		2	371	311	83.82
311	293	94.21		3	311	293	94.21		3	311	311	100.00
364	357	98.08		4	364	364	100.00		4	364	359	98.62
426	391	91.78		5	426	413	96.95		5	426	360	84.62
1754	1628	92.82		Total	1754	1611	91.85		Total	1754	1623	92.53

	Class	Pixels	Cor_Pix	% Corr		Class	Pixels	Cor_Pix	% Corr
	1	400	400	100.00		1	5044	5044	100.00
	2	162	153	94.44		2	212	212	100.00
	3	409	405	99.02		3	283	283	100.00
	4	118	118	100.00		4	1226	1226	100.00
	5	367	367	100.00		5	250	250	100.00
	Total	1456	1443	99.11		Total	7015	7015	100.00
	MATHER	TRAD			SEED	UNREF			
	Class	Pixels	Cor_Pix	% Corr		Class	Pixels	Cor_Pix	% Corr
	1	282	282	100.00		1	282	282	100.00
	2	371	311	311.00		2	371		70.88
	3	311	311	311.00		3	311		97.74
	4	364	360	360.00		4	364	364	100.00
	5	426	362	362.00		5	426		96.94
	Total	1754	1646	92.70		Total	1754		92.70

Table 3.15 Classwise Classification accuracy with different techniques on the Singrauli Image.

Plate 3.1 **The growth of training area by seed pixel method in the marshy region of Bhopal Image, the white region indicates the seed pixel specified by the user.**

This plate emphasizes that in a heterogeneous classes like marshy area the conventional technique of training area delineation is tedious and may be unrepresentative too.

Legend

Colour	Class
Black	Water
Red	Forest / vegetation
Green	Marsh
Blue	Urban / Residential area
Pink	Scrub

Plate 3.2 Classified image of Bhopal using traditional training area.

Legend

Colour	Class
Black	Water
Red	Forest / vegetation
Green	Marsh
Blue	Urban / Residential area
Pink	Scrub

Plate 3.3 Classified image of Bhopal using training area generated by seed pixel algorithm.

Legend

Colour	Class
Black	Water
Red	Forest / vegetation
Green	Marsh
Blue	Urban / Residential area
Pink	Scrub

Plate 3.5 **Classified image of Bhopal using traditional training area, with Buttner's refinement technique applied on it at k=2.**

The classification results are much better visually than at k=1.

Legend

Colour	Class
Black	Reservoir Water
Red	Forest / vegetation
Green	Mine Dumps
Blue	Rocky Terrain
Pink	Water Bodies

Plate 3.6 Classified image of Singrauli using training area generated by seed pixel algorithm.

Legend

Colour	Class
Black	Reservoir Water
Green	Forest / vegetation
Red	Mine Dumps
Grey	Rocky Terrain
Blue	Water Bodies

Plate 3.7 **Classified image of Singrauli using traditional training area, with Buttner's refinement technique applied on it at k=1.**

Chapter 4

CONCLUSIONS AND SCOPE FOR THE FUTURE WORK

4.1 Conclusions

Training area selection and refinement is considered to be an important part of the supervised classification procedure, as it affects the overall classification accuracy of the classifier. Obtaining a representative class accurately and efficiently has been the major goal of this study, which has been attempted by the refinement of traditional data and by applying alternate methods of training area selection..

The impetus for the development of non-traditional approach to training field extraction grew out of the increasing evidence of the inadequacy and inefficiency of traditional training approaches when applied to higher resolution data. A novel technique, termed the seed pixel method, of training area extraction was envisioned as a means for improving both the efficiency and consistency of training field extraction for images with high resolution and high spectral variability. It was designed to increase the degree of automation of training field extraction, thereby significantly reducing analyst time requirements, while maintaining acceptable levels of accuracy. It is viewed as a potential replacement to other non-traditional image analysis techniques.

The basic approach used in the thesis was to apply two levels of classification strategies, one by classifying the images by selecting training areas by conventional methods and by seed pixel algorithm method. And the second, by classifying the images after refinement suggested by Buttner et al. (1988) and Mather (1987). The classified images generated from these methods are then compared to the test areas (delineated on the basis of topographic maps, published by the Survey of India). Finally, the accuracies of these classifiers have been compared by generating confusion matrices.

On the basis of the above experiments, the following conclusions may be drawn:

A. The Seed Pixel approach to training area extraction

1. Accuracy

The following trends were observed for various classes and overall accuracies:

(a) Training area classification accuracy using all seed pixel approach is always 100%. It means that the training area selected by this algorithm are highly statistically separable, which is demonstrated by the divergence analysis of the training area too.

(b) The 100% training area accuracy does not translate into 100% test area accuracy by the seed pixel method. However, the performance of the ML classifier using training areas from this method is comparable to the results of classification using training areas from traditional refinement procedures.

It is concluded that, by the proper selection of thresholds and window size, by more experiments, we can expect even better test area classification results.

(c) The application of refinement techniques on the training area obtained by the seed pixel approach, created problems of getting zero covariance matrix for some classes. Although, the problem was tackled by retaining a few pixels in the data set, so that the covariance matrix did not become singular.

(d) The traditional training areas selected by this method gave high overall classification accuracies. It means that most of the classes were highly separable. The training area selected by the seed pixel method also gave good results. But, the point to be noted is, it worked very well where traditional methods failed. In both the study sites, the traditional technique gave low classification accuracy for natural heterogeneous classes i.e. forests and rocky terrain. The seed pixel algorithm increased the classification accuracy of these classes. Hence it can be concluded that the seed pixel method is specially useful for classification in heterogeneous environments.

2. Efficiency

The time requirements for the training data extraction and analysis phases of the procedures are markedly reduced with the non-traditional approach for both the study sites. The relative simplicity of implementing the non-traditional approach was a further advantage in this context. Use of the new procedures was

straightforward, requiring very little prior knowledge of the techniques for parameter specification and permitting substantial reductions in the number of repetitive decisions required of the analyst. Features of irregular shapes, and those having holes in it can also be selected by this method, which is not possible by traditional methods.

B. Refinement Procedures

1. Statistics upgrading technique

This technique proposed by Mather (1987), gave marginal improvements compared to the accuracies obtained for the unrefined data. Since, the unrefined data accuracies very high, more increase in accuracy by statistics upgrading is not very probable. This will need more investigations to prove its utility for training data refinement.

2. Pixel Removal Technique

This method proposed by Buttner et al. (1988) showed marked changes in the training area statistics after refinement. The main observations from the experiments performed are:

1. The divergence values increase by application of this technique. This means better separability between classes, less confusion and hence better accuracy of classification.

2. The method gave improved results at k=2, when the number of pixels retained is sufficient enough to just characterize the whole class.

It can be concluded that this technique gives marked improvement in classification accuracy, which is supported by the observation of increasing divergence with refinement. The k=2 value gave better results for most cases, hence this value is highly recommended for refinement.

Hence it can be said that the seed pixel algorithm is an efficient and excellent method for improved truing area extraction. The Buttner's technique of refinement gives better accuracy results than the unrefined data. It is strongly recommended these methods be used in classification procedure and be incorporated in image processing systems.

4.2 Future Directions of Work

More tests on images of different cover types are recommended to gain better understanding of the refinement procedures. This method should be experimented upon data of different spatial resolution to check its effectiveness. If the normality condition is followed in the data set or not, remains to be checked.

Although not specifically demonstrated in this study, one of the biggest advantage of seed pixel algorithm may be its potential for integration with existing GIS processing capabilities and data-bases. Seed pixel coordinates, and conceivably even certain data extraction parameters, could be derived from existing GIS data layers (e.g., the size and broad cover categories of potential training areas could be extracted from GIS data and

used to set maximum field size and threshold parameters for the seed pixel algorithm). Using the GIS input, the seed pixel approach could be used for general land cover classification for an entire region or to statistically characterize spectral information at specific locations within a region for change detection studies.

References

Arai K. (1992). Supervised thematic mapper classification with a purification of training samples, *International Journal of Remote Sensing*, Vol.13, 2039-2049.

Buttner G, Hajos T and Korandi M. (1989). Improvements to the Effectiveness of Supervised Training Procedures, *International Journal of Remote Sensing*, Vol.10, 1005-1013.

Buchheim P.M. and Lillesand T.M. (1989). Semi-Automated Training Field Extraction and Analysis for Efficient Digital Image Classification. *Photogrammetric Engineering and Remote Sensing*, Vol.55,No.9, September 1989, pp. 1347-1355.

Campbell, N.A. (1980) : Robust procedures in multivariate analysis.I : Robust covariance estimation. *Applied Statistics*, 29, 231-237.

Dikshit, O. (1992) *The Classification of Texture in Remotely Sensed Environmental Imagery*, Ph.D. Thesis, University of Cambridge.

Gill G. and Gill G. (1994). A Simple Cleaning Procedure for Improvement of Training Site Statistics. *Photonirvachak, Journal of Indian Society of Remote Sensing*, Vol.22, No.2, 1994.

Hand, D. H. (1981) *Discrimination and Classifciation*, John Wiley and Sons, Chichester.

Jensen, J. R. (1986) *Introductory Digital Image Processing: A Remote Sensing Perspective*, Prentice-Hall, Eaglewood Cliffs, New Jersey.

Lillesand, T. M. and Keifer, R. W. (1987) *Remote Sensing And Image Interpretation,* 2nd ed., John Wiley and Sons, New York.

Mather, P. M. (1987) *Computer processing of remotely-sensed images,* John Wiley and Sons.

Mather, P.M. (1987). Preprocessing of Training Data for Multispectral Image Classification, *Proceedings of the Annual Conference of the Remote Sensing Society, University of Nottingham*, september 7-11, 111-120pp

Mather, P. M. (1990) "Theoretical problems in image classification", *Steven, M. D. and Clarke, J. A. (eds.), Application of Remote Sensing in Agriculture*, Buttnerworth, London.

Singh, A. (1984) "Some clarification about the pairwise divergence measure in remote sensing", *International Journal of Remote Sensing*, Vol. 5, No. 3, pp. 623-627.

Swain, P. H. and Davis, S. M. (1978) *Remote Sensing: the quantitative approach*, Newyork McGraw-Hill.